Sonora: Inclusive Voice Play For Children With Various Abilities

Daniil Umanski
Holon Institute of Technology
Golomb 52, Holon, Israel
made@hit.ac.il

Yael Avni
Holon Institute of Technology
Golomb 52, Holon, Israel
yaela@hit.ac.il

Michal Rinott
Shenkar College of Enginering,
Design and Art
Ramat Gan, Israel
michalrin@gmail.com

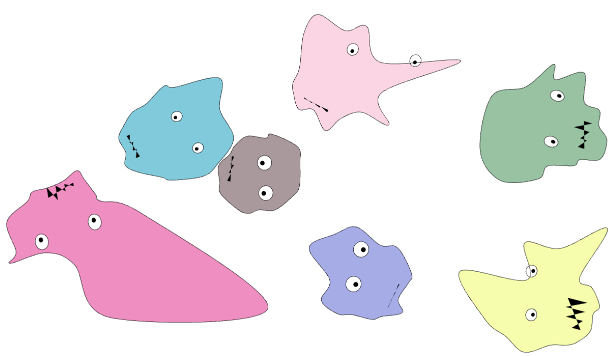

Figure 1: Sonora screen-shot

CHI-PLAY'17, Octorber 15 – 18, 2017, Amsyerdam.
ACM XXX-X-XXXX-XXXX-X/XX/XX.

Abstract
Sonora is an exploratory interactive installation that explores human voice through a playful transition between the physical and digital worlds. Sonora is designed to elicit expressive, naturalistic human vocal sounds, and to explore the full range of capability of the human voice through playful and inclusive interaction. This work-in-progress paper describes the conception, interaction design process, technical details, and planned elaborations of the project.

Author Keywords
Interaction Design; Vocal Exploration, Inclusive Design

ACM Classification Keywords
H.5.m [Information interfaces and presentation]: User interfaces, Voice I/O.

Introduction
Our voice is an integral part of our self-experience, of our communication and self-expression. In this paper we present and discuss an ongoing work with placing human voice in the center of a playful interaction. The project is a continuation of our previous work focusing on play, children and self-expression[13]. Our goal is to empower voice-body connections in children through expressive, joyful vocal exploration, as well as to make this exploration inclusive for children with a range of vocal, motor, and cognitive abilities. To that end,

we develop and evaluate Sonora – a computer game designed to stimulate children to experiment with their voice. We outline some of the opportunities and challenges in designing a game which affords prolonged engagement, curiosity, and motivation throughout play, as well as discuss its potential merits in a therapeutical context.

The experience

Players are invited to push a large button and start voicing into a microphone. In the world of the game, their voice is visualized in real-time as a new voice-creature, reflecting the phonetic characteristics of the voice. The newly created voice-creature then falls into a virtual world, inhabited with other voice-creatures. Players witness their vocalizations come to life, as voice-creatures receive real-world physical properties, propel in motion and collide with voiced shapes of other players in dynamic virtual world. While voice-creatures constitute a homogeneous population, they retain enough individual character to surprise and delight. The projected world is in constant motion, collisions between creatures trigger layers of sound output, comprised of the original voices of players, creating a humorous, surprising soundscape.

Design Principles

With Sonora, we aimed to design a game which will sustainably engage children with exploring their voice, lead them to experiment again and again with the various ways of shaping their voice-creature. The principle design themes are playfulness, engagement and the embodied nature of voice interaction [8, 4]. Simplicity is another important design value in the project, and one single button constitutes the entire game interface.

Figure 2: Voice experimenting with Sonora.

Inclusive Design

Our approach draws strongly from the ideas of inclusive and universal design, with the aim of crafting an interaction which will allow as wide as possible range of users to participate [15]. Because voicing is not constrained by motor or speech abilities, vocal experimentation presents an inclusive opportunity for self-expression. While traditional forms of voice biofeedback systems tend to expect a 'correct' production, Sonora equally welcomes any vocalization, opening the play experience for a spectrum of vocal, motor and cognitive abilities.

Embodiment

Contrary to screen-based interactions, Sonora affords a more personal, embodied play. The transformation from voice to voice-creatures resonates with the concept of *dupliances*, which are defined as devices that encompass both physical and virtual activity, as opposite to purely information based appliances [5]. In Sonora, this is accomplished by combining the physical aspect of vocal play with the virtual dynamic world where voices are captured and animated [3].

Voice-Body connection
In the field of music therapy, working with voice is practiced to create voice-body relations, as well as to evoke positive emotions. Singing or holding a tone of voice while exploring emotions or parts of the body are used to empower people to utilize their resources, similar to community musicking[2]. Music therapist Kenneth Bruscia writes that: "Being an inner instrument of the body, the voice is at a unique and powerful vantage point for working with the self from within. The voice is powerful and yet vulnerable since it is constantly in connection with our body through breathing and the bloodstream. The voice is something we always bring with us. It is also vulnerable because it reveals a person's emotions and expresses her identity."

Implementation

Interface Object
The tangible interface object consists of a box, where a large button and a microphone are embedded. The microphone utilizes the mental model of microphones being an invitation for voicing [12]. The large red button (echoing the ubiquitous Record button) functions to elicit curiosity, and allows even children with low motor abilities to engage in interaction. Upon button press, the system is opened for audio input, and remains 'listening' until it detects a silence of 500 milliseconds, thereby closing audio input and generating the newly created vocal creature into the game world.

Audio Processing and Visualization
We process incoming audio in real-time within Max/MSP, utilizing Tristan Jehan's analysis externals, which provide estimations of loudness (time-domain power energy), pitch (voice fundamental frequency), and Bark-scale spectral decomposition (perceptually modelled spectrum bands). Analysis data is relayed to a Processing sketch, where the visuals are generated. Upon button press, a new voice-creature

Figure 3: The interface object.

is generated as a round blob with 24 equally-spaced control points. Then, throughout the voicing phrase, the shape undergoes a continuous transformation driven by the incoming audio. Real-time analysis data are accumulated, resulting in a unique visual signature of the vocalization as a whole.

The geometric transformation is based on mapping the amplitudes of Bark-scale spectral bands to the blob's control points. In other words, the creatures' body acts as a dynamic 'equalizer' of the voice signal. The overall size of the creature is determined by the loudness of the vocalization. Furthermore, the parameter of voice noisiness (bark-based spectral flatness measure) is mapped to the creatures restitution parameter (the bounciness of the body in the simulated world). Together, the physical characteristics of the voice-creature (size, shape, restitution) serve to individualize the newly added voice and affect its behaviour in the virtual world.

Usability testing

In the process of developing Sonora, we have tested the game for usability with 8 children, aged 3-6. The insights from these trials allowed us to consider key interaction design questions:

Record button

We tested a design whereby recording is activated while the button is kept pressed. This proved to be difficult for the younger children, and expected to be even more so for children with compromised motor abilities. Therefore, in the subsequent design we require only a single press to activate the recording, while its termination is done automatically by detecting silence longer than 500 milliseconds. This tends to work better, however, other alternatives must be tested in order to arrive at the desired usability. For example, automatic voice onset detection could trigger the recording, making the button redundant altogether. While this option may seem attractive in the sense of offering a 'zero-interface' gameplay, the interface object itself is an attractive entry point to the game. In our usability trials, children were fascinated by the object and the sense of direct touch they could exercise with it.

Visual feedback

In the current version of Sonora, upon button press, a new voice-creature of a default size appears in the game. Surprisingly, for some children this alone was satisfying enough to keep pressing the button to create more creatures, with no urge to transform the creatures with their voice. In order to increase the motivation to voice, we plan to test alternative designs whereby the default creature size ceases to be rewarding, and warrants an active voicing from the player in order to grow.

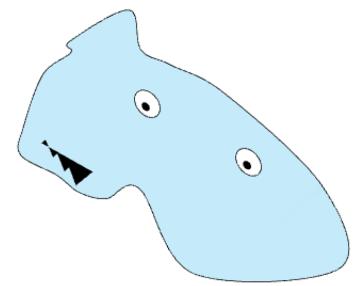

Figure 4: A voice-creature.

Discussion

Visualizing voice

An intriguing design challenge central to the game is a mapping of auditory to visual parameters which is both intuitive as well as being able to elicit prolonged experimentation with that mapping. One the one hand, there seems to be a need in elaborate enough visual variation between various voicing explorations in order to stimulate and empower exploration. On the other hand, if the mapping relies on non-intuitive or not perceptible auditory parameters, the visualization might be perceived as arbitrary and interaction will weaken. Research on optimal audio visualization strategies is only emerging [7]. In a recent study it was reported that representing both pitch and loudness combined in a single visualization is preferred over visualization in only one dimension[10]. In our current implementation we try to balance between visualizing straightforward parameters (amplitude of main spectral bands) and the less obvious features (noisiness and brightness) as secondary parameters. However, we choose bark-scale spectral bands over pitch since pitch tracking in explorative, naturalistic sounds is often irregular and brisk.

Speech-Language Therapy
In order to evaluate the potential therapeutic merits of Sonora, we have conducted a survey with 20 practising Speech-Language therapists. Each was given a 5 minute play session with the game, and subsequently answered questions pertaining to the clinical utility of the game. The survey results revealed that therapists regard Sonora as a highly potential tool to elicit vocal productions and train vocal skills in the following clinical contexts:

- Stimulation of vocalizations for children with minimal vocal output, such as in dysarthria, apraxia of speech, cerebral palsy, autism.
- Production of appropriate vocal intensity, for children who cannot normalize vocal intensity, and tend to speak with overly low or hight volume.
- Stimulation of vocal production for children who do not speak at all due to developmental delay or selective mutism.
- Improve control of vocalizations for children with hearing impairments, as a means of learning to map inner vocal sensations to produced voice through visual feedback.

Most intervention for voice disorders in children aim to raise the child's awareness of his or her own voice production [14]. Specific aims may address breathing patterns, habitual pitch and loudness, or aiming to decrease tension and vocal effort when speaking[11]. In this context, Sonora offers an engaging method for stimulating vocal exploration and expending the dynamic vocal range.

Future Work

Therapeutic perspective
From a therapeutic perspective, initial trials and the survey conducted with speech therapists are encouraging, as there exists a significant challenge in prompting non-verbal children to vocalize and establish a relationship with their voice. In that regard, Sonora holds a strong promise to become a valuable tool. Beyond vocal stimulation, granting children with special needs a sense of control in gameplay may strongly improve their sense of self-reliance and self-esteem. Furthermore, inclusively designed play may promote the sense of belonging, as the child takes part in an activity that is also performed by typical peers and siblings. In that sense, the play experience may become a bridge for social inclusion. We therefore plan to further explore the therapeutic potentials of the Sonora project.

Joint Voicing
A fascinating perspective we envision in the development of the game is designing a joint-voicing interaction, whereby two players create voice-creatures through voicing together. In such setup, each player uses a separate microphone, while 2 incoming audio signals are analyzed in parallel. The setup provides a rich ground for exploring, promoting and testing dynamics of vocal synchrony and coordination. The design space is significant, ranging from more practically oriented dynamics of modelling desired voice patterns (one of the players being a clinician), turn-taking and imitation, to more ambient interactions, whereby vocal entrainment and resonance can promote empathy and connectedness between players.

Acknowledgements

We kindly thank all the children, parents and therapists who participated in our trails.

REFERENCES

1. Elena Bodrova and Deborah J Leong. 2003. The Importance of Being Playful. *Educational Leadership* 60, 7 (2003), 50–53.

2. Birgitta Cappelen and Anders-Petter Andersson. 2016. Health Improving Multi-Sensorial and Musical Environments. In *Proceedings of the Audio Mostly 2016*. ACM, 178–185.

3. Inger Ekman and Michal Rinott. 2010. Using vocal sketching for designing sonic interactions. In *Proceedings of the 8th ACM conference on designing interactive systems*. ACM, 123–131.

4. Johan Fagerlönn, Stefan Lindberg, Anna Sirkka, and Gunnar Oledal. 2016. OREGANO: Play a concert organ using your voice. In *Proceedings of the Audio Mostly 2016*. ACM, 229–233.

5. Daniel Fallman, Niklas Andersson, and Lars Johansson. 2001. Come together, right now, over me: Conceptual and tangible design of pleasurable dupliances for children. In *1st International Conference on Affective Human Factors Design, Singapore*.

6. Gordon Muir Giles and Barbara E Chandler. 1997. *The essence of play: a child's occupation*. Amer Occupational Therapy Assn.

7. Joshua Hailpern, Karrie Karahalios, Laura DeThorne, and Jim Halle. 2010. Vocsyl: Visualizing syllable production for children with ASD and speech delays. In *Proceedings of the 12th international ACM SIGACCESS conference on Computers and accessibility*. ACM, 297–298.

8. Daniel Hug, Moritz Kemper, Kathleen Panitz, and Karmen Franinović. 2016. Sonic Playgrounds: Exploring Principles and Tools for Outdoor Sonic Interaction. In *Proceedings of the Audio Mostly 2016*. ACM, 139–146.

9. Amit Pitaru. 2008. E is for everyone: The case for inclusive game design. *The ecology of games: Connecting youth, games, and learning* (2008), 67–86.

10. Rebecca S Schaefer, Lilian J Beijer, Wiel Seuskens, Toni CM Rietveld, and Makiko Sadakata. 2016. Intuitive visualizations of pitch and loudness in speech. *Psychonomic bulletin & review* 23, 2 (2016), 548–555.

11. Ingo R Titze. 2006. Voice training and therapy with a semi-occluded vocal tract: rationale and scientific underpinnings. *Journal of Speech, Language, and Hearing Research* 49, 2 (2006), 448–459.

12. Nicholas True, Nigel Papworth, Ru Zarin, Jeroen Peeters, Fredrik Nilbrink, Kent Lindbergh, Daniel Fällman, and Anders Lind. 2013. The voice harvester: an interactive installation. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. ACM, 3003–3006.

13. Daniil Umanski and Michal Rinott. 2016. The Drawbox Project. *Proceedings of the Conference on Computation, Communication, Aesthetics and X.* (2016).

14. Catherine Y Wan, Loes Bazen, Rebecca Baars, Amanda Libenson, Lauryn Zipse, Jennifer Zuk, Andrea Norton, and Gottfried Schlaug. 2011. Auditory-motor mapping training as an intervention to facilitate speech output in non-verbal children with autism: a proof of concept study. *PloS one* 6, 9 (2011), e25505.

15. Jacob O Wobbrock, Shaun K Kane, Krzysztof Z Gajos, Susumu Harada, and Jon Froehlich. 2011. Ability-based design: Concept, principles and examples. *ACM Transactions on Accessible Computing (TACCESS)* 3, 3 (2011), 9.